**CURSO: CC471**

Practica 2.

**1. Introducción**

PRACTICA: ALINEAMIENTO DE SECUENCIAS.

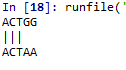
1. **Alineamiento de secuencias con el método de Dot – Plot**

1.1 Dadas dos secuencias de nucleótidos:

A = ACTGG

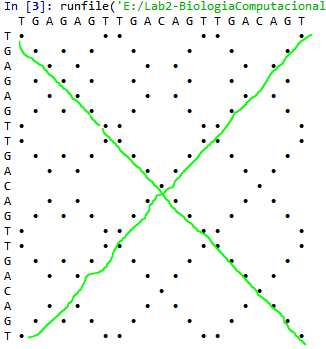
B = ACTAA

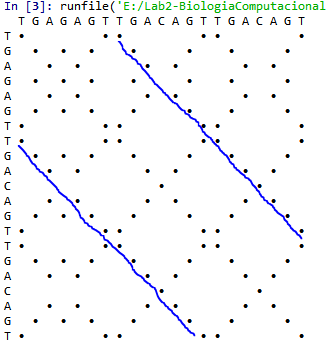
Realizar el dotplot correspondiente - haga un screenshot.

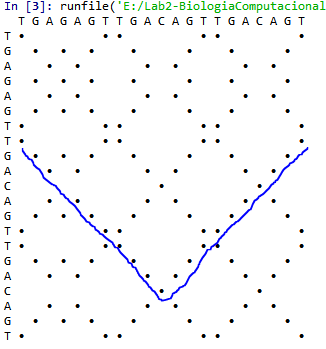


* 1. Dada una secuencia con repeticiones en **tandem perfectas**. Un **dotplot** obtenido a partir de la comparación de esta secuencia consigo misma:

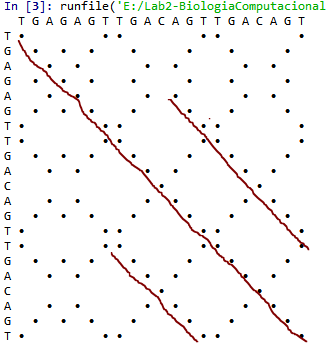
Utilizando la secuencia: TGAGAGTTGACAGTTGACAGT

a. Muestra dos diagonales que se cruzan en el centro (X).

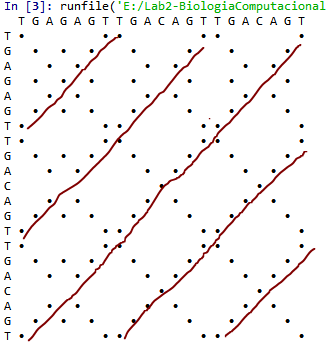
b. Muestra dos diagonales paralelas

c. Muestra dos diagonales invertidas que se cruzan en un vértice (V).

d. Muestra una diagonal principal y varias diagonales más pequeñas correspondientes

a las repeticiones.

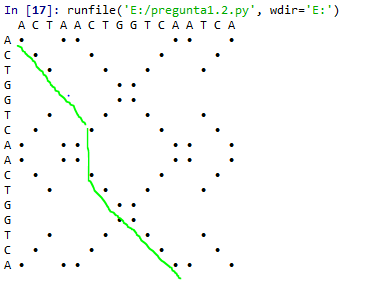
e. Muestra una diagonal principal y varias diagonales más pequeñas invertidas

correspondientes a las repeticiones.

1.3 Dibuje el dotplot para las siguientes secuencias y deduzca (sin usar el algoritmo) cuál es el camino que tomaría un algoritmo de alineamiento global:

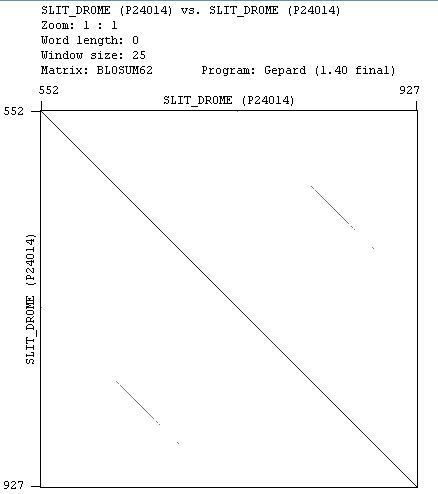
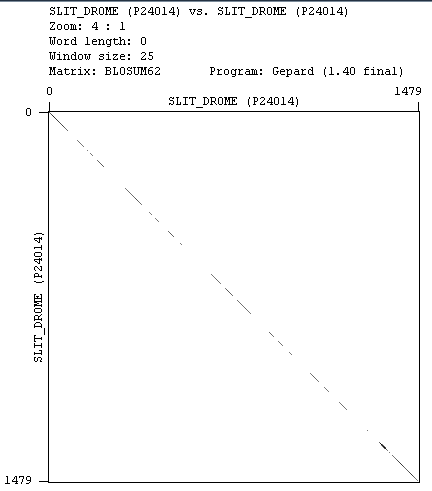
A = ACTGGTCAACTGGTCA

B = ACTAACTGGTCAATCA

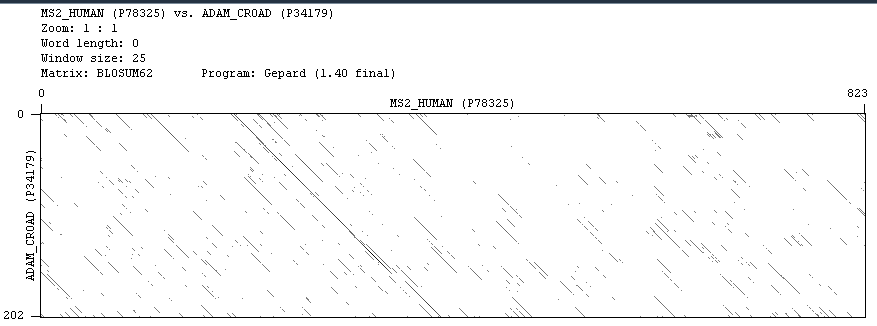


1.4 Realizar un dotplot de la siguiente secuencia consigo misma. ¿Qué características tiene esta secuencia? Puede usar GEPARD (http://cube.univie.ac.at/gepard)

Cuando se cambia las dimensiones del plot, se genera **diagonales paralelas**.

******Escala 1479 Escala 927**

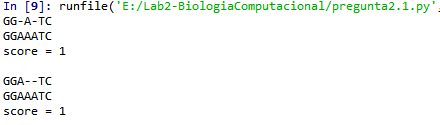
1.5 Dadas las siguientes secuencias, realizar un dotplot. ¿Tienen algo en común? Realizar un alineamiento local de ambas secuencias.

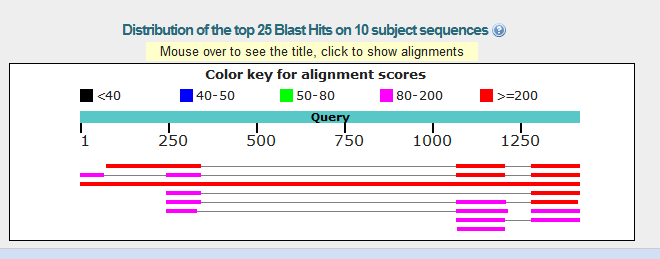
Posee **muchas diagonales paralelas** a la diagonal principal

**2. (Realizar Alineamientos)**

2.1 Usando los parámetros por defecto para match (1), mismatch (-1) y gap (-2), calcule el score para cada uno de los siguientes alineamientos:

GG-A-TC GGA--TC

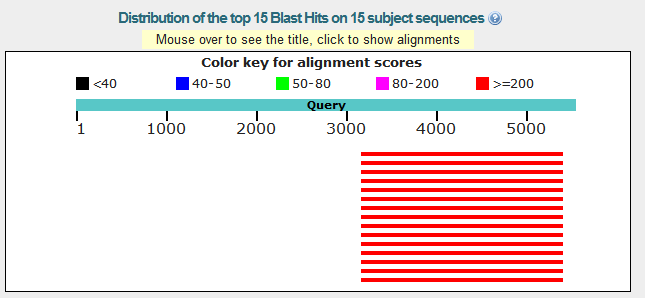
GGAAATC GGAAATC

2.2 Busque en alguna base de datos biológicas dos secuencias de proteínas que tengan la misma función, p.ej homologas a la proteina L36. P. Ej podria usar la base de datos Protein del sitio de NCBI para buscar 3 homologos. p. Ej (en los organismos Homo Sapiens, mus musculus, y otro organismo que usted elija)

**Homo Sapiens**

Realice 2 alineamientos globales (Homo Sapiens - Mus Musculus y Homo Sapiens – el

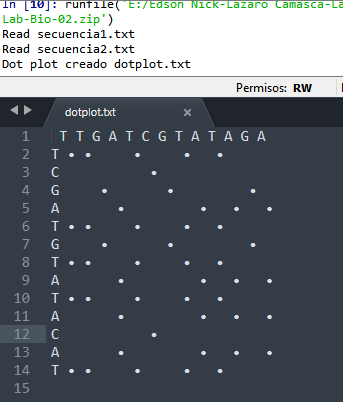
organismo que Usted escogió) utilizando el formato FASTA de cada una de las secuencias de las proteinas. Para esto puede utilizar el algoritmo BLASTP de NCBI

****  **Mus Musculus**

Grabe los resultados obtenidos y las secuencias en formato FASTA utilizadas

**3. (Codificar)**

1. Modificar el programa que hizo para tomar las secuencias biológicas desde un archivo.

2. Modificar el programa que hizo para poder escribir el dotplot a un archivo de nombre “salida.txt”.